Defining Treatment Effects: 
A Regulatory Perspective 
Some Thoughts on Implementation of E9(R1)

Thomas Permutt
FDA Topic Leader
Three Questions and an Answer

• Words or numbers?

• Treatment policy
  – Are you going to let them stop doing it?
  – Are you going to make us do it?

• Principal stratification—Seriously?

• Robust and easy per-protocol analyses
Words or Numbers?

• Framework—Words
• No new methods—No numbers
• Can we agree on statistical methods first, and write about estimands later?
Words or Numbers?

• Can we agree on statistical methods first, and write about estimands later?

• No.

• The framework is important, but ...

• We also want different methods

• Words *and* numbers
  – Hypothetical ≠ MAR
Treatment Policy

• Are you going to let them stop doing it?
• Are you going to make us do it?
• “Outcome” studies—Keep doing it
• “Symptom” studies
  – “In symptomatic settings, it is not the usual practice to continue to assess effectiveness in patients after they have stopped taking the assigned treatment (ITT approach)” (Temple & O’Neill 2012)
Which Way?

Current: All patients but with imputation/MAR

NO

Redefine ITT

yes

Treatment policy

yes

Other strategies
Treatment Policy (“Symptom”)

• Are you going to let them stop doing it?
  – They never started

• Are you going to make us do it?
  – You have to stop pretending to do it
  – Do it or ...
  – Do something else and say what
Principal Stratification

• Impractical?
• Irrelevant?
Impractical?

• Hard to understand
• Hard to satisfy assumptions
• Hard to prespecify
Hard to Understand?

• Not really

• Part of *treatment effect* is to make subjects continue or discontinue
  – {continue, discontinue} × treatment → principal strata

• If you can’t tell me what you did with subjects who would discontinue only on test drug, I won’t understand
  – I.e., principal stratification is an essential part of the *discussion*, even when not of the solution
Hard to satisfy assumptions

• Yes, really
• Sometimes easier than MAR
Antarctica

• Some subjects move for reasons completely unrelated to treatment
• MCAR, so ...
• Can use completers
Antarctica

• Some subjects move for reasons completely unrelated to treatment
• MCAR, so ...
• Can use completers
• Right?
Antarctica

- Some subjects move for reasons completely unrelated to treatment
- MCAR, so ...
- Can use completers
- Right?
  - No, not necessarily MCAR
  - Yes, can use completers ...
  - To estimate effect in principal stratum
Hard to prespecify

• Yes
• Consider prespecifying modeling algorithm rather than model?
• With cross-validation
• But maybe it is impractical
  – Maybe selection modeling is not better than outcome modeling
    • Or maybe it is
    • Or maybe do both (double robust)
  – But modeling is modeling
    • Looking hard is a feature, because it is hard
    • Don’t redefine ITT
Irrelevant?

• Want
  – Pharmacologic effect or ...
  – Per-protocol effect or ...
  – “Efficacy” (vs. “effectiveness”)

• This is difficult to define
  – Part of treatment effect is to make subjects continue or discontinue

• Principal stratification (uniquely?) can yield precise definitions

• It is hard, maybe impractical
  – Can see it’s hard (good!)
  – Easy ways are not easy
Robust Per-Protocol Analyses

• Not analyses of per-protocol set
• Crosscountry method (Permutt and Li 2017)
• Undilution method (Permutt and Hebel 1989)
Not Per-Protocol Set

• Don’t estimate population variance by sample variance
  – Because it’s *biased*

• Don’t estimate treatment effect by difference in means
  – Use ANCOVA
  – Because it’s less *variable*
  – But estimates same estimand

• So don’t estimate per-protocol effect by per-protocol dataset!
Crosscountry Scoring

• Start 7, count best 5
• If your (test) 5 beat my (placebo) 5, your team is faster
• Nothing assumed about other 2, they just don’t count
But ...

• Inefficient
  – Not very, even compared to no dropouts
    • Think about median
  – Not comparable to imputation
• Unfair
  – Not.
• Not clinically meaningful
  – Sometimes, but ...
  – Are you sure the raw mean is more meaningful?
  – If the worst scores are important, you’d better get them
Undilution

• Assume ...
  – All treatment effect due to taking active drug
    • No compliance effect in controls
    • No persistent effect in noncompliers
  – No one in control group takes active drug
  – (Sensitivity analysis needed)

• Results
  – Half of active group comply
  – Treatment policy effect is 5

• What is effect in compliers?
Treatment Policy Dilutes

• TP = P{comply} * {complier effect} + P{not comply} * {noncomplier effect}
• 5 = (0.5) * X + (0.5) * 0
• X = 10
• Undilute!
Robust Per-Protocol Analyses

• Exist
• Do not need to solve hard problem
• But don’t use a bad solution to the hard problem instead of a robust solution to an easy problem
Summary

• Treatment policy
  – Yes or no
  – Not redefined

• Principal stratification
  – Maybe too hard
  – But you can see how hard it is
  – Therefore better than hard methods that look easy

• Robust per-protocol analyses are possible
  – If you don’t try to do the hard problem